AD_____

Award Number: DAMD17-98-1-8019

TITLE: Development of General Methods for Identification of Genes Regulated by Oncogenic Stimuli or DNA Damage

PRINCIPAL INVESTIGATOR: Stephen Elledge, Ph.D.

CONTRACTING ORGANIZATION: Baylor College of Medicine
Houston, Texas 77030

REPORT DATE: May 2001

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

20010727 086

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	May 2001	Annual (13 Apr 00 - 12 Apr 01)

4. TITLE AND SUBTITLE
Development of General Methods for Identification of Genes Regulated by Oncogenic Stimuli or DNA Damage

5. FUNDING NUMBERS
DAMD17-98-1-8019

6. AUTHOR(S)
Stephen Elledge, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Baylor College of Medicine
Houston, Texas 77030

E-Mail: selledge@bcm.tmc.edu

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT ***(Maximum 200 Words)***

In order for cells to prevent mutations, they have an elaborate regulatory network that is capable of sensing and responding to chromosomal aberrations. When damage to chromosomes is detected, cells induce a set of genes that respond to the damage in order to repair the damage and to prevent the cell cycle from progressing. In addition, cells have set up another elaborate regulatory network to ensure that the key events of the cell cycle occur in the correct order with the correct timing. Both the DNA damage response and the cell cycle regulatory pathways are critical to prevent the types of events that generate cancers. Furthermore, both of these pathways operate by the precise regulation of transcription of genes involved in this process. Our proposal has centered around identification of DNA damage inducible genes and genes important in the cell cycle by either retroviral insertions or, more recently by microarray hybridization. We have now identified several hundred genes regulated by these events, either DNA damge or the cell cycle and have learned a great deal about the processes the cell employes to prevent the accumulation of cancer-causing mutations.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Breast Cancer	19
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

Table of Contents

(4) Introduction

The ability of cancer cells to proliferate inapproriately is due the loss of tumor suppressor genes and the gain of function of oncogenes. Two classes of tumor suppressor genes exist, mutators and growth regulators. Mutators are genes that when mutant cause an increase in the rate of genomic instability and hence the rapid accumulation of mutations of all classes. Genes of the mutator class include mis-match repair genes involved in HPCC, DNA repair genes such as X.P., cell cycle checkpoint genes such as ATM and p53. Recently two breast cancer genes BRCA1 and BRCA2 have been implicated in DNA repair and fall into this class. The second class of tumor suppressors, growth regulators, are those directly involved in regulating cell proliferation or the ability of tumor cells to survive and metastasize. Oncogenes such as myc or ras are dominant and act in opposition to tumor suppressors. Many oncogenes and tumor suppressors regulate the cell cycle and affect celll cycle regulation. In order to understand this regulation, it is important to have knowledge of genes under cell cycle transcriptional regulation. Mitosis represents one of the most fundamental activities of eukaryotic cells, playing an essential role in events ranging from gametogenesis and wound-healing to multicellular development[1-3]. Proper regulation of gene activity during the cell cycle is therefore likely to govern critical aspects of diverse biological processes. Indeed, disruption of cell cycle-dependent mRNA regulation results in profound phenotypic consequences. For example, defects in the retinoblastoma (Rb) and E2F-1 proteins, which control transcription at the G_1/S transition, precipitate developmental defects, aberrant cell morphologies, and uncontrolled proliferation[4-6]. The pleiotropic effects of these mutations suggest that a considerable proportion of all genes experience differential regulation during the cell cycle. However, although these transcripts have been comprehensively catalogued in yeast, few targets of cell cycle-dependent transcriptional machinery have been identified in higher organisms[7,8]. A global perspective on genetic regulation during eukaryotic mitosis might therefore significantly expand our understanding of both normal cell division and the abnormal phenotypes observed in certain pathological processes.

In addition,, the transcriptional induction of genes in response to DNA damage is a critical component of the cellular response to DNA damage and the prevention of mutagenesis. The identification of transcriptional target genes is critical to understanding how these regulatory pathways function. In this grant we propose the identification of target genes such as these to better our understanding of these processes and to provide tools for future analysis.

(5) Body

As noted in our previous reports, during the process of devising retroviral vectors to identify genes under the control of promoters regulated by DNA damage and other stimuli, we began a collaboration with several groups to identify genes induced by these stimuli using transcriptional arrays. Our retrovial work was never able to identify more than just a few genes that were reproducibly induced by DNA damage. Thus due to the massive amount of work involved in identifying the gene for each insertion, we opted to perform transcriptional profiling using microarrays. This report details our success in this area.

Transcriptional Profiling of Cell Division and DNA Damage

In this study, we measured cell cycle regulated transcripts in human cells on a genome-wide scale using high-density oligonucleotide arrays. Primary fibroblasts prepared from human foreskin were grown to approximately 30% confluence and synchronized in late G_1 using a double thymidine block protocol[9]. Cultures were released from arrest, and cells were collected every two hours for twenty-four hours, covering nearly two complete

cell cycles. Messenger RNA was isolated, labelled, and hybridized to arrays containing probes for approximately 40,000 human genes and non-overlapping ESTs[10]. The entire experiment was performed in duplicate under identical conditions for a subset of these genes for an independent confirmation of results.

The two cell cycle data sets were concatenated and analyzed using both supervised and unsupervised clustering of expression patterns (See Methods). Five groups of transcripts were identified that displayed a periodicity consistent with the length of the cell cycle in both data sets, with peaks corresponding to early G_1, late G_1, S, G_2, and M phase. Using this approach, 1,332 transcripts, including 895 previously uncharacterized ESTs, were assigned to cell cycle-regulated expression clusters. These assignments were made only for transcripts that clearly oscillated in obvious patterns, and the conservative nature of this analysis makes it possible that some cell cycle regulated transcripts were not selected. Consistent with previous studies, multiple cyclin family members were identified, as well as *Cdc2*, *Cdc8*, *Cdc25* and numerous genes involved in DNA replication and chromatin structure[4]. This information can be viewed and downloaded at the Internet address www.mag.com/expressionanalysis.

We also sought to better understand the relationship between transcriptional changes that normally occur during mitosis and those observed during events that affect the progression of cell division. For the purposes of this assessment, transcript levels were quantitated in fibroblasts after exposure to ultraviolet light and methyl methane sulfonate (MMS), both of which precipitate DNA damage and consequent cell cycle arrest at multiple checkpoints.

All experiments were performed in duplicate and compared with untreated cells to identify consistent differences. Stringent requirements were also used to identify these differences. To qualify, changes had to be called in at least two independent experiments and more than a two-fold difference was required. Of the more than 6,000 genes assayed, 64 were consistently induced and 24 repressed after UV exposure, and 73 were induced and 18 repressed after MMS exposure. These data are also available at the Internet site given above.

Systematic Functional Analysis of Expression Clusters

Previous large-scale expression studies have relied heavily on subjective evaluations of the functions of differentially expressed genes[7,11,12]. However, the rapidly increasing number of these data sets has created an urgent need for more standardized methods for assessing the coordinate regulation of biological pathways. Overrepresentation of genes from a common functional class within an expression cluster may reveal global pathway activation[13]. We implemented a statistical analysis based on this hypothesis. More than 2400 characterized human genes represented on the arrays were classified into 160 functional categories (Methods). Using the binomial distribution function, *P* values were calculated based on the frequency of occurrence of genes with common biological function in each expression cluster. Statistically significant enrichments for specific functions were detected in four of the five cycle-regulated transcription clusters, with *P* values ranging from 1.4×10^{-4} to 1.9×10^{-11}.

Transcriptional Regulation of Central Mitotic Processes

The most basic test of the systematic analysis applied in this study is the detection of biological processes known to play a central role in cell division. As anticipated, we found coordinate upregulation of transcripts involved in cell cycle control, DNA replication,

chromosome segregation, and cytokinesis, providing an important validation of this approach.

The comprehensive identification of specific genes that experience cell cycle-dependent transcriptional control in these known pathways may provide insight into how regulatory disruption may exert biological effects. We first examined transcripts for genes that govern cell cycle progression ($P = 4.3 \times 10^{-12}$) Cyclins such as *A*, *B*, and *F* family members were coordinately induced at the G_2-M transition, at which point they regulate the entry into mitosis. However, some genes that inhibit the G_1-S transition (such as p16*INK4* and *Myt1*) were induced at this time period, while others (such as p27*Kip1*) were maximally repressed. The most striking functional coordination was observed for genes involved in proteolytic control of cell cycle progression, including *Cdc5*, *Cdc20*, *Cks1*, *Cdc34*, *E2A*, and the cyclin-specific ubiquitin ligase *Ubc10*, all of which are induced during M phase. Notably, a second set of ubiquitin-associated transcripts was highly expressed during the G_1-S transition, including the Nedd4-like ubiquitin protein ligase *Wwp2*, ubiquitin-specific protease *Usp7*, and cullin *Cul4A*.

Consistent with findings in yeast, numerous DNA replication genes were induced in late G_1 and S phase, including some genes known to be positively regulated by E2F transcription factors, such as DNA polymerase alpha, DNA primase, components of replication factor A and C, MCM genes, histones, and dihydrofolate reductase ($P = 4.4 \times 10^{-5}$) (Table 2). Interestingly, we also observed unexpected overrepresentation of genes related to actomyosin function in the S phase cluster ($P = 6.7 \times 10^{-4}$). Many of these transcripts encode proteins directly implicated in cytokinesis, including several smooth muscle myosins, alpha-actin, the small GTPase Rho, gelsolin Sm22, calponin Cnn1, and Tau. Induction during S phase was also observed for numerous effectors of actin-based cytoskeletal remodeling not previously linked to cell division; for example, the regulatory genes *Arf6*, *Cdk5*, *Tiam1*, and *Hef1*. Other transcripts found in this cluster encode gelsolin-like actin polymerization domains (such as *Fli1* and *Capg*), but have not been assigned a more specific biological role.

Genes that regulate chromosome segregation were coordinately upregulated during M phase, including the spindle checkpoint sensor *Mps1*, the centrosomal kinase *Nek2*, and stathmin, a putative control point for microtubule assembly ($P = 8.4 \times 10^{-10}$)[14-16]. Nearly global upregulation during M phase was also observed for microtubule-dependent motor proteins that physically effect chromosome segregation, such as mitosin and virtually every kinesin known to interact directly with the centromere[17,18]. Clear cell cycle regulation was not observed for dynein genes, suggesting restriction of this transcriptional control to a subset of motor proteins.

Cell Cycle Control of Extracellular Matrix Activity and Apoptosis

These analyses reveal that an unexpectedly high proportion of the genes induced during G_2 phase function in the extracellular matrix ($P = 1.9 \times 10^{-11}$). Genes found in this cluster encode cell-cell adhesion molecules such as collagen VI alpha-1, integrin alpha-6, desmoplakin, plakoglobin, and laminin. Other matrix-associated genes found in both the G_2 and M clusters positively regulate cell movement through the extracellular matrix; for example, platelet-derived growth factor receptor *PdgfA*, vasodilator-stimulated phosphoprotein *Vasp*, hyaluron synthase *Has2*, hyaluron receptor *Rhamm*, tissue plasminogen activator *Plat*, and urokinase plasminogen activator receptor[19-22].

Marked enrichment was also detected in the late S phase cluster for Bcl-related transcripts ($P = 2.1 \times 10^{-4}$), which have been extensively characterized as inhibitors of apoptotic proteases[23]. These genes encode Bcl-7A, Bcl-xL, Bak, Mcl-1, the Bcl-2

interacting proteins Bbc3 and Bad, and the Bcl-xl interacting protein Bap31. Interestingly, no other group of apoptotic inhibitors were overrepresented in any of the cell cycle expression clusters. It was also observed that none of the Fas or TNF-related genes that positively regulate apoptosis were found to show a significant degree of cell cycle regulation in these data.

DNA Damage Response Involves S phase Transcripts

It is possible that cell cycle regulated transcripts comprise organized programs of biological activities that are required outside of cell division. DNA damage in *Saccharomyces cerevisiae*, for example, is known to induce a significant proportion of the transcripts induced during late G_1 phase (M. Campbell and S. Tavazoie, unpublished data). Such overlapping specificities may identify cellular responses that involve cell cycle-dependent transcriptional machinery.

In the 6,800 transcripts surveyed in this study, 42 were reproducibly induced in response to both UV light and MMS in human fibroblasts. This group included many known damage-induced transcripts such as *c-FOS*, *JUNB*, *GADD45*, and *IEX1*. Strikingly, these genes were found to share considerable target specificity with transcripts in the S phase cluster ($P = 2.3 \times 10^{-8}$) but not with any other group of cell cycle regulated genes. FACS analyses indicate that these cells were not arrested in S phase when samples were collected, excluding the possibility that such an effect was responsible for the observed induction of these transcripts.

Yeast genes induced by DNA damage

Since the goal of this project was to identify DNA damage inducible genes from eukaryotes, we thought it would be useful to also know the spectrum of DNA damage-inducible genes in a model eukaryote S. cerevisiae. We have established a collaboration with Pat Brown's lab to establish the complete list of DNA damage inducible genes from yeast using chip hybridization technology[9]. We have performed a number of analyses and have identified approximately 300 transcripts that are significantly increased in response to DNA damage. Many of these have obvious human homologs that we can test for DNA damage-inducibility. Interestingly we found that they fall into several categories based on time of induction during a time course and into groups of genes based on function. These genes are shown grouped into functional units in Tables 1 and 2 below. In addition to genes involved in DNA replication, nucleotide synthesis and DNA repair we were surprised to see a number of genes involved in electron transport, carbohydrate metabolism, Amino acid biosynthesis, protein folding and degradation. This data has now been submitted for publication.

Table 1

Types of yeast genes induced by DNA damage by MMS through a 2 hr time course.

Total number of genes induced over 2-fold	number of genes peaking at 15-30 min	number of genes peaking at 45-60 min	number of genes peaking at 90-120 min
Amino Acid Biosynthesis			
22	14	4	4
Protein Degradation			
39	6	13	20
Protein Folding			
7	6	1	0
Transport			
20	7	4	9
Nucleotide synthesis			
15	0	1	14
Stress-induced genes			
20	3	2	15
Electron transport			
8	0	2	6
Carbohydrate metabolism			
45	0	7	38
DNA replication and repair			
13	0	6	7

Wild type cells were treated with 0.03%MMS (80% viability)

Table 2. Names of 4 classes of yeast genes induced by DNA damage

DNA replication and/or repair	**Nucleotide Biosynthesis**
DUN1 Checkpoint protein kinase	THI1
RNR1* Ribonucleotide reductase	THI5
RNR2* Ribonucleotide reductase	THI12
RNR4* Ribonucleotide reductase	TRR1
TRX2*	TRR2
MAG1 Methyl transferase	UGP1
RFA1 DNA Replication	YNK1
RFA2 DNA Replication	ADE17
POL30 DNA Replication	STF2
PHR1 Photo-repair	YBR284W (similar to AMP deaminase)
RAD51 Recombination	RNR1*
TRR1	RNR2*
YHR029C *(similar to thymidylate synthase)	RNR4*
	TRX2*
	YHR029C *(similar to thymidylate synthase)

Stress-induced genes	**Genes involved in protein degradation**
SOD1 Superoxide dismutase	RPN6
HYR1	RPT4
YAP1 Transcription factor	RPT6
YHB1	RPN7
TIR1	RPT2
XBP1 Transcription factor	RPN12
TIR2	RPT1
HSP42 Heat shock protein	RPN3
HSP104 Heat shock protein	RPT3
HSP82 Heat shock protein	RPT5
HSP12 Heat shock protein	PRE5
HSP26 Heat shock protein	PRE9
SSE2 Heat shock protein	PRE10
STI1 Heat shock protein	PUP3
SIS1 Chaperone	PRE3
YGP1	PRE2
USV1	PRE8
DDR48	PUP1
GRE2	PRE1
GRE3	UMP1
	UBA1
	UBC5
	UBC8
	UBC13
	UBI4
	UFD1
	YPS6
	YPS6
	APE2
	PRB1
	PEP4

PRD1
LAP4
CDC48
UBP6
YAP180
AUT7
APG1
APG7

Discussion

The expression-based analyses used in this study revealed both known and unexpected results. DNA replication genes were strongly upregulated during late G_1, as were numerous cell cycle control and chromosome segregation genes during M phase (Table 1). However, the observed global upregulation of genes related to actomyosin function and programmed cell death during S phase, and induction of extracellular matrix-associated genes during G_2, indicates that a much broader set of biological pathways is transcriptionally activated during cell division than was previously known. Notably, whereas eight biological functional classes were found to be strongly overrepresented in cell cycle regulated expression clusters, none were significantly underrepresented. One interpretation of this observation is that acquisition of cell cycle-specific regulatory elements occurs stochastically in the genome, but is retained by genes for which mitotic regulation confers a biological advantage.

These data offer new insights into established mechanisms of cell cycle regulation. For example, the coordinate induction of ubiquitin-associated cell cycle control genes during M phase indicates a requirement for restricting the time of activity of these genes, and may reveal how aberrant transcription directly interferes with proteolytic cell cycle regulation. Interestingly, these results are in direct contrast to observations made in *Saccharomyces cerevisiae*, where transcriptional control of proteolytic regulators is observed only rarely[7]. Recent evidence has implicated cullin family members in the proteolytic degradation of G_1-S regulators, supporting the possibility that the ubiquitin-associated genes induced at this time may participate in checkpoint regulation[24].

These data also implicate previously characterized genes in specific mitotic activities. For example, the nucleotide excision repair genes *Rad2* and *Ercc1* are induced during S phase, suggesting that they may function in the correction of misincorporation in newly synthesized DNA. We also observed that the DNA damage-responsive gene *Gadd45*, increased levels of which have been shown to block passage through a G_2-M checkpoint, is transcriptionally repressed during this period[25]. Naturally low levels of *Gadd45* transcript during G_2-M may permit cell cycle progression in the absence of DNA damage-dependent induction, providing a sensitive means of checkpoint regulation[26].

Examination of the actomyosin-related transcripts observed to be induced during S phase reveals that many of these genes govern assembly of the contractile ring. Rho-mediated reorganization of actin stress fibers is essential for cytokinesis, and its deregulation has been implicated in tumor-specific abnormalities in cell separation[27]. Calponin contains protein domains that are required in yeast for actin binding during cytokinesis, while the bacterial homolog of Tau has been shown to maintain integrity of the contractile ring[28,29]. Coordinate transcriptional upregulation of this set of genes occurs unexpectedly early in the cell cycle, well before the onset of mitosis.

A number of the actin-related transcripts induced during S phase are not directly implicated in cytokinesis. Their gene products may be responsible for more global cytoskeletal changes that accompany mitosis. For example, Arf6 directs reorganization of cortical actin at both endosomes and the plasma membrane and has recently been implicated in cell migratory activities[30,31]. Both Cdk5 and Tiam1 modulate Rac activity and are required for the proper development of actin-based outgrowths such as neurites[32]. Other genes such as *Fli1* and *Capg* encode gelsolin-like actin polymerization domains and are hypothesized to function in cellular motility, but have not been assigned a specific role in cell division[33]. These results suggest that diverse cytoskeletal remodeling pathways are activated to facilitate the gross morphological changes required during cell division.

The observed cell cycle regulation of these numerous cytoskeleton-related signaling targets offers an opportunity to examine the temporal organization of their transcriptional behavior. The induction in S phase of *Rho* with regulatory and downstream effector genes such as Rho GDP dissociation inhibitor, actin, and *Tau*, and the simultaneous expression of *Rac* with genes such as *Arf6*, *Cdk5*, and *Tiam1* suggest that genes in the same signaling pathway are generally co-regulated. We examined cell cycle regulated genes functionally related to another small Ras-like GTPase, Cdc42, which is known to participate in signaling cascades for actin and filamin reorganization. Specifically, Ack kinase and components of the Arp2/3 actin polymerization complex have been linked to signaling pathways involving Cdc42, but not Rho or Rac[34,35]. Interestingly, we found that transcripts encoding all of these proteins were co-induced during G_2 phase. Such temporal compartmentalization could increase discrimination between closely related signaling pathways. It is also possible that phase of induction separates cytoskeletal remodeling programs that must be activated sequentially during mitosis.

The systematic analysis applied here is perhaps most valuable in detecting unexpected coordinate regulation of functional pathways, which is likely to prove difficult to assess based solely on subjective surveys of genome-scale experimental data. We identified an overrepresentation of transcripts with extracellular matrix-associated functions in the G_2 cluster. At least some of these gene products, such as plakoglobin, desmoplakin, integrin alpha-6, and laminin, represent components of structures such as desmosomes and hemidesmosomes, which facilitate intercellular adhesion and communication[36]. Transcriptional upregulation of these genes likely prepares dividing cells for re-establishment of contact and communication with the extracellular environment. However, a far greater number of matrix-associated transcripts in the G_2 and M clusters are known to positively regulate cellular motility, in some cases through signaling to the actin cytoskeleton. This group of genes includes *HAS2*, *RHAMM*, *VASP*, *PLAT*, and urokinase plasminogen activator. It is possible that these genes play a major role in the separation and migration of daughter cells following mitosis. Interestingly, these gene products have been characterized most extensively in terms of their ability to increase the motility of tumor cells. Overexpression of either Has2 or Rhamm is sufficient to promote development of invasive cellular phenotypes[20,22]. Therefore, restriction of the activity of these genes to defined cell cycle intervals may represent a general mechanism for enabling activites required for proliferation while limiting tumorigenic cellular behavior.

The global repression of Bcl-2 related antiapoptotic genes throughout G_2 and M is consistent with findings that deactivation of this particular class of genes is partially responsible for an increased susceptibility to programmed cell death during this period[37]. The induction of antiapoptotic genes during S phase may also serve to counteract the proapoptotic effects of the E2F-1 transcription factor, which positively regulates the G_1/S transition[38]. Recent evidence has supported an unexpectedly broad role for Bcl-related genes in cell cycle progression, possibly through direct regulation of the E2F class of

transcription factors or other activities downstream of cyclin-dependent kinase 2 (Cdk2)[39-41]. One candidate effector for cell cycle regulation of *Bcl*-related genes is the Stat5 transcription factor, which is known to positively regulate transcription from the *Bcl-xL* promoter and whose expression we also observed to be restricted to S phase[42]. Further experiments will be required to evaluate the full functional significance of these observations.

The significant overlap between DNA-damage induced and S phase transcripts indicates that common subsets of genes may be required for response to apparently unrelated cellular conditions. This phenomenon has been observed in *S. cerevisiae*, where widely divergent biological processes have been linked by their dependence on the proper function of a common set of genes[43]. One explanation might be that numerous genes that function directly in DNA damage repair are required for the proper execution of chromosome replication. However, none of the transcripts that appear in both groups participate directly in DNA synthesis. Instead, these genes function predominantly in the stress response, either through intracellular signaling (the transcription factors *Jun-B*, *Atf3*, *Atf4*, *Egr1*, and *Egr2* and the dual specificity phosphatases *Dusp1* and *Dusp5*) or extracellular communication (ligands including *IL-6*, *IL-11*, and *EGF*-like growth factor). Therefore, it appears likely that the general regulatory response to cellular stress is governed partially by transcriptional mechanisms that also coordinate gene activity during the cell cycle. It remains to be determined what other transcriptional responses display significant similarities with the expression clusters identified from these data.

A number of genes that are clearly cell cycle-regulated, including *Edn-1*, *Cox-2*, *Tfpi-2*, *Plat*, *Pai2*, and *Plaur*, were previously reported to be transcriptionally activated effectors of the serum response[12]. Because one effect of serum deprivation is mitotic synchronization, the previously observed transcriptional results are unlikely be specific to the serum response. Further analysis and comparison with the previous results will allow identification of serum-dependent transcriptional effects that are not simply the result of cell cycle arrest. These data should provide a helpful general reference point for future human profiling studies that involve aspects of cell cycle progression.

More than 800 transcripts from previously uncharacterized EST sequences were also found to be clearly cell cycle regulated. Because many human open-reading frames have no significant similarity to characterized genes, prediction of biological function based entirely on sequence homology is limited. Identification of transcriptional behavior under different conditions provides an additional means for assessing the function of these genes. Approximately 60% of the previously characterized genes identified in this study as cell cycle regulated were known to play a mitotic phase-specific role, compared to 5% of all genes represented on the arrays. Therefore, observation of cell cycle regulation for a transcript indicates a more than 10-fold increase in the probability of mitotic function.

The cell cycle cluster to which an uncharacterized EST is assigned provides additional clues to its biological role. For example, based on the calculations made about biological function in this paper, genes upregulated during M phase have a more than 20-fold higher likelihood than random genes of participating in cell to cell communication. Furthermore, assessing these transcriptional results in the context of sequence homology may provide additional clues to cellular role[11]. For example, one uncharacterized EST induced during G_2/M shows similarity to katanin, an ATPase whose oligomerization leads to microtubule severing during cell division[44,45]. Another EST, induced in S phase, displays high amino acid identity with the *RTEF1* transcription factor, which preferentially induces muscle-specific genes[46,47]. This gene may therefore be partially responsible for the global cell cycle regulation of actin and myosin transcripts. The increasing number of genome-scale transcriptional data sets will soon generate multiple lines of evidence about

the cellular role of many newly sequenced or previously uncharacterized genes. These data should provide opportunities for more systematic comparisons between expression and DNA sequence data for the purposes of predicting gene function.

Although the perspective provided by this study is restricted to the transcriptional level, the detection of functions ranging from DNA replication to cytoskeletal reorganization indicate that these data compose a representative survey of mitotic activities in higher eukaryotes. The global regulation of processes such as extracellular matrix reorganization and inhibition of apoptosis suggests that cell cycle-dependent transcriptional control influences the execution of many biological programs not obviously related to mitosis. In marked contrast to the limited changes observed following deletion of specific transcription factors[26,48], the regulatory fluctuations observed here and in profiling of other general cellular processes involve broad functional classes of genes and multiple members of gene families[12,49]. Analogous to quantitative genetic phenomena, certain biological activities may depend on the aggregation of transcriptional effects whose significance is difficult to measure on an individual basis.

The functional classifications applied in this study represent only a first attempt to formally structure public domain knowledge about human genes for the purpose of systematically analyzing experimental genomic data. Generation of increasingly refined representations of such information should enable more a far more sophisticated understanding of this and other transcriptional data sets. Certainly, many biological activities may be mediated through a limited number of transcriptional targets and will therefore not meet the statistical thresholds inherent to this approach. However, given the considerable size and proliferation of expression data sets, functional evaluations now represent the rate-limiting step to deriving meaning from these studies. The methods demonstrated here provide a first solution to this growing problem. Because of the scope of this data set, more conclusions remain to be drawn than can be described in this manuscript. We encourage the reader to directly access the data via the Internet or FTP.

As for the yeast DNA damage inducible genes, we have found a large number of genes that were inducible in a Mec1 and Dun1-dependnet manner. Several sets of these inducible genes are represented by general oxidative stress or by protein unfolding. We observed a large number of proteins involved in protein folding as well. These will serve as a point of comparison for the mammalian genes isolated now and in the future to determine the similarities between organismal responses to DNA damage.

Methods

Synchronization of Cells

Human cell lines used in this study were derived from human foreskin samples[50]. Cells were grown in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum and antibiotics (all from GIBCO). Cells were grown in 150 cm^2 flasks to 25-30% confluence and synchronized using a double thymidine block protocol[9], with a first block for 16 h, a 10 h release, and a second block for 15 h. The final concentration of thymidine used in the block medium was 2 mM. Approximately 2 x 10^7 cells were harvested at each time point. About 10% of cells from each time point were fixed in ethanol and stained with propidium iodide for flow cytometry analysis, and the remainder were immediately frozen at -80 °C for subsequent isolation of RNA.

UV and MMS treatment

Cells were grown asynchronously and passaged before reaching 80% confluency. At the last passage, cells were combined and split equally to 18 150 cm^2 dishes and grown overnight (~16 h). Media was then removed from each dish and stored separately. All dishes were then rinsed once with 1X PBS. Six dishes were treated with UV irradiation (40 J/m^2) and six dishes were treated with MMS at a final concentration of 0.015% (v/v). The remaining six dishes were not subjected to any treatment. Media was then added back to all dishes and all samples were incubated another 4 h before harvest. For harvest, cells were trypsinized for removal from plates, immediately flash frozen in liquid nitrogen, and stored at –80 °C .

Isolation and Hybridization of mRNA

Total RNA was isolated and biotin-labeled as previously described[10]. Samples were hybridized to a total of five arrays containing approximately 1.4 x 10^6 oligonucleotide probes to approximately 40,000 human genes and ESTs (Affymetrix, Santa Clara). Hybridizations were carried out at 42 °C for 14 to 16 h with mixing on a rotisserie at 60 rpm. Following hybridization, the solutions were removed, the arrays were rinsed with 6X SSPE-T (0.9 M NaCl, 60 mM NaH_2PO_4, 6 mM EDTA, 0.005% Triton X-100 adjusted to pH 7.6), rinsed with 0.5X SSPE-T (75 mM NaCl, 5 mM NaH_2PO_4, 0.5 mM EDTA, 0.005% Triton X-100 adjusted to pH 7.6), and incubated with 0.5X SSPE-T at 42 °C for 15 minutes. Following washing, hybridized biotinylated cRNA was fluorescently labeled by incubation with 2 μg/ml streptavidin-phycoerythrin (Molecular Probes, Eugene, OR) and 1 mg/ml acetylated BSA (Sigma, St. Louis, MO) in 6X SSPE-T at 42 °C for 10 min. Unbound streptavidin-phycoerythrin was removed by rinsing at room temperature prior to scanning. The arrays were read at a resolution of 3 μm using a specially designed confocal scanner (Affymetrix, Santa Clara, CA) as described previously[10].

Expression Data Analysis

Data were normalized between samples by setting the mean hybridization signal for each sample equal to 1000 arbitrary units and the standard deviation to 1500 arbitrary units after discarding the top 2% and the bottom 2% of all signals. The hybridization signal for each gene or EST was then normalized to a unit standard deviation with a mean of zero. Genes having similar intensity patterns across multiple time points were detected by clustering using the Pearson correlation coefficient as the metric of similarity. Genes were placed into specific clusters by identifying relevant expression patterns and matching the expression profile of each gene to a pattern using a correlation coefficient threshold.

The abundance of cell cycle phase-specific transcripts is expected to fluctuate periodically. We identified periodic patterns using supervised and unsupervised clustering methods. Supervised clustering was based on comparisons to series of sine waves offset by consecutive time intervals, similar to previous analyses of cell cycle expression data[8]. Typically, we demanded that a gene be correlated to one of the sine wave patterns by a Pearson's correlation coefficient of at least 0.70 over 24 samples. However, many known cell cycle-regulated transcripts displayed patterns that did not conform to simple sine waves. Therefore, we also searched for common patterns in the data by conducting unsupervised clustering between all genes, searching for groups of at least four transcripts that exhibited a correlation coefficient of at least 0.90. Twenty-two such groups were identified, of which five were found to be cell cycle-regulated. These groups were then averaged and used as seed patterns for clustering against all genes, using a correlation coefficient of at least 0.70. Groupings from supervised and unsupervised clustering with identical peak times were

merged, resulting in the clusters described in the text. More than 90% of these transcripts were scored as detectable in more than a quarter of the timepoints.

A database of functional information for more than 2400 human genes was generated by manual placement of genes into functional categories using information from Medline and other publicly available information sources. The database contains 160 controlled-vocabulary categories that are classified hierarchically. The database schema is similar to that of the MIPS yeast database (www.mips.org), but also contains functions relevant to multicellular organisms. This database is commercially available through Molecular Applications Group (www.mag.com).

The binomial distribution function was applied using the Stingray Expression Analysis software package (Molecular Applications Group, Palo Alto, CA). Briefly, if the number of genes in an expression cluster with a given biological function was less than or equal to the number expected by random chance, then the *P* value was calculated using the formula:

$$P=\left[\begin{matrix}n\\x\end{matrix}\right]p^{x}(1-p)^{n-x}$$

where n is the number of genes in a given cluster, x is the observed number of genes found in both a given cluster and a given functional category, and p = the overall frequency at which genes of a given functional category are found in the genes represented on arrays. If the number of genes in an expression cluster with a given biological function was greater than the number expected by random chance, then the *P* value was calculated using the formula:

$$P=1-\sum_{i=}^{x-}\left[\begin{matrix}n\\i\end{matrix}\right]p^{1}(1-p)^{n-1}$$

Table 1 Human genes expressed maximally in G2/M.

Accession Number		Gene function
X66945_at	3	H.sapiens N-sam mRNA for fibroblast growth factor receptor
X54942_at	3	H.sapiens ckshs2 mRNA for Cks1 protein homologue
X79201_at	3	H.sapiens mRNA for SYT
Z36714_at	3	H.sapiens mRNA for cyclin F
X87613_at	3	H.sapiens mRNA for skeletal muscle abundant protein
Z49989_at	3	H.sapiens mRNA for smoothelin
L25876_at	3	Homo sapiens protein tyrosine phosphatase (CIP2)mRNA,
M62994_at	3	Homo sapiens thyroid autoantigen (truncated actin-binding protein)
U47635_at	3	Human D13S824E locus mRNA, complete cds
X14850_at	3	Human H2A.X mRNA encoding histone H2A.X
U14518_at	3	Human centromere protein-A (CENP-A) mRNA, complete cds
M25753_at	3	Human cyclin B mRNA, 3' end
U73379_at	3	Human cyclin-selective ubiquitin carrier protein mRNA,
M96803_at	3	Human general beta-spectrin (SPTBN1) mRNA, complete cds

Accession Number		Gene function
M86699_at	3	Human kinase (TTK) mRNA, complete cds
U37426_at	3	Human kinesin-like spindle protein HKSP (HKSP) mRNA,
D26361_at	3	Human mRNA for KIAA0042 gene, complete cds
D31885_at	3	Human mRNA for KIAA0069 gene, partial cds
D43948_at	3	Human mRNA for KIAA0097 gene, complete cds
X51688_at	3	Human mRNA for cyclin A
D78514_at	3	Human mRNA for ubiquitin-conjugating enzyme, complete cds
U63743_at	3	Human mitotic centromere-associated kinesin mRNA, complete cds.
U28386_at	3	Human nuclear localization sequence receptor hSRP1alpha mRNA,
U05340_at	3	Human p55CDC mRNA, APC activator
U01038_at	3	Human pLK mRNA, complete cds
M90657_at	3	Human tumor antigen (L6) mRNA, complete cds
M91670_at	3	Human ubiquitin carrier protein (E2-EPF) mRNA, complete cds

Table 2. Human genes maximally expressed during S phase.

Accession Number		Gene function
X68277_at	7	H.sapiens CL 100 mRNA for protein tyrosine phosphatase
X66087_at	7	H.sapiens a-myb mRNA
X89750_at	7	H.sapiens mRNA for TGIF protein
Z46967_at	7	H.sapiens mRNA for calicin (partial)
X62535_at	7	H.sapiens mRNA for diacylglycerol kinase
Z24725_at	7	H.sapiens mitogen inducible gene mig-2, complete CDS
M77140_at	7	H.sapiens pro-galanin mRNA, 3' end
Z24727_at	7	H.sapiens tropomyosin isoform mRNA, complete CDS
U62015_at	7	Homo sapiens Cyr61 mRNA, complete cds
Z46629_at	7	Homo sapiens SOX9 mRNA
J02854_at	7	Human 20-kDa myosin light chain (MLC-2) mRNA, complete cds
U73960_at	7	Human ADP-ribosylation factor-like protein 4 mRNA, complete cds
U44975_at	7	Human DNA-binding protein CPBP (CPBP) mRNA, partial cds
U27655_at	7	Human RGP3 mRNA, complete cds
U27768_at	7	Human RGP4 mRNA, complete cds
U59752_at	7	Human Sec7p-like protein mRNA, partial cds
U28049_at	7	Human TBX2 (TXB2) mRNA, complete cds
X16416_at	7	Human c-abl mRNA encoding p150 protein
M92934_at	7	Human connective tissue growth factor, complete cds
U04636_at	7	Human cyclooxygenase-2 (hCox-2) gene, complete cds
U18300_at	7	Human damage-specific DNA binding protein p48 subunit (DDB2)
U81607_at	7	Human gravin mRNA, complete cds
X52142_at	7	Human mRNA for CTP synthetase (EC 6.3.4.2)
D90209_at	7	Human mRNA for DNA binding protein TAXREB67
X52599_at	7	Human mRNA for beta nerve growth factor
X58377_at	7	Human mRNa for adipogenesis inhibitory factor
L08246_at	7	Human myeloid cell differentiation protein (MCL1) mRNA
U08021_at	7	Human nicotinamide N-methyltransferase (NNMT) mRNA,
U40490_at	7	Human nicotinamide nucleotide transhydrogenase mRNA,
M12174_at	7	Human ras-related rho mRNA (clone 6), partial cds
M62831_at	7	Human transcription factor ETR101 mRNA, complete cds
J03764_at	7	Human, plasminogen activator inhibitor-1 gene, exons 2 to 9

(6) Key Research Accomplishments

A) Identification of all known DNA damage inducible genes in the yeast S. cerevisiae. Described in previous report

B) Determination of cell cycle regulated genes in human diploid fibroblasts, 40,000 transcripts analyzed.

C) Determination of DNA damage-inducible genes in human diploid fibroblasts, 6,800 transcripts analyzed.

(7) Reportable Outcomes.

Paper submitted to Molecular Biology of the Cell, in review. A second paper published in Nature Genetics

Cho, R.J., Huang, M., Campbell, M.J., Dong, H., Steinmetz, L., Sapinoso, L., Hampton, G., Elledge, S.J., Davis, R.W and Lockhart, D.J. (2001) Transcriptional regulation and function during the human cell cycle. *Nat. Genet.* 27: 48-54.

(8) Conclusions

So far our work has established that a large number of human genes are controlled in a cell cycle regulated manner. Many processes previously not known to be cell cycle regulated were found to be regulated in this manner. Furthermore, many cell cycle regulated genes are also inducible by DNA damage.

(9) References

1. Nicklas, R.B. How cells get the right chromosomes. *Science* **275**, 632-7 (1997).
2. Nasmyth, K. Viewpoint: putting the cell cycle in order. *Science* **274**, 1643-5 (1996).
3. Stillman, B. Cell cycle control of DNA replication. *Science* **274**, 1659-64 (1996).
4. Sanchez, I. & Dynlacht, B.D. Transcriptional control of the cell cycle. *Curr Opin Cell Biol* **8**, 318-24 (1996).
5. Dyson, N. The regulation of E2F by pRB-family proteins. *Genes Dev* **12**, 2245-62 (1998).
6. Friend, S.H. *et al.* A human DNA segment with properties of the gene that predisposes to retinoblastoma and osteosarcoma. *Nature* **323**, 643-6 (1986).
7. Cho, R.J. *et al.* A genome-wide transcriptional analysis of the mitotic cell cycle. *Mol Cell* **2**, 65-73 (1998).
8. Spellman, P.T. *et al.* Comprehensive identification of cell cycle-regulated genes of the yeast Saccharomyces cerevisiae by microarray hybridization. *Mol Biol Cell* **9**, 3273-97 (1998).
9. Rao, P.N. & Johnson, R.T. Mammalian cell fusion: studies on the regulation of DNA synthesis and mitosis. *Nature* **225**, 159-164 (1970).
10. Lockhart, D.J. *et al.* Expression monitoring by hybridization to high-density oligonucleotide arrays [see comments]. *Nat Biotechnol* **14**, 1675-80 (1996).
11. Chu, S. *et al.* The transcriptional program of sporulation in budding yeast [published erratum appears in Science 1998 Nov 20;282(5393):1421]. *Science* **282**, 699-705 (1998).
12. Iyer, V.R. *et al.* The transcriptional program in the response of human fibroblasts to serum [see comments]. *Science* **283**, 83-7 (1999).

13. Tavazoie, S., Hughes, J.D., Campbell, M.J., Cho, R.J. & Church, G.M. Systematic determination of genetic network architecture [see comments]. *Nat Genet* **22**, 281-5 (1999).
14. Hardwick, K.G., Weiss, E., Luca, F.C., Winey, M. & Murray, A.W. Activation of the budding yeast spindle assembly checkpoint without mitotic spindle disruption. *Science* **273**, 953-6 (1996).
15. Fry, A.M., Arnaud, L. & Nigg, E.A. Activity of the human centrosomal kinase, Nek2, depends on an unusual leucine zipper dimerization motif. *J Biol Chem* **274**, 16304-10 (1999).
16. McNally, F.J. Microtubule dynamics: Controlling split ends. *Curr Biol* **9**, R274-6 (1999).
17. Chan, G.K., Schaar, B.T. & Yen, T.J. Characterization of the kinetochore binding domain of CENP-E reveals interactions with the kinetochore proteins CENP-F and hBUBR1. *J Cell Biol* **143**, 49-63 (1998).
18. Schaar, B.T., Chan, G.K., Maddox, P., Salmon, E.D. & Yen, T.J. CENP-E function at kinetochores is essential for chromosome alignment. *J Cell Biol* **139**, 1373-82 (1997).
19. Liu, K. *et al.* Reversible tumorigenesis induced by deficiency of vasodilator-stimulated phosphoprotein. *Mol Cell Biol* **19**, 3696-703 (1999).
20. Kosaki, R., Watanabe, K. & Yamaguchi, Y. Overproduction of hyaluronan by expression of the hyaluronan synthase Has2 enhances anchorage-independent growth and tumorigenicity. *Cancer Res* **59**, 1141-5 (1999).
21. Entwistle, J., Hall, C.L. & Turley, E.A. HA receptors: regulators of signalling to the cytoskeleton. *J Cell Biochem* **61**, 569-77 (1996).
22. Nguyen, D.H., Hussaini, I.M. & Gonias, S.L. Binding of urokinase-type plasminogen activator to its receptor in MCF- 7 cells activates extracellular signal-regulated kinase 1 and 2 which is required for increased cellular motility. *J Biol Chem* **273**, 8502-7 (1998).
23. Adams, J.M. & Cory, S. The Bcl-2 protein family: arbiters of cell survival. *Science* **281**, 1322-6 (1998).
24. Feng, H. *et al.* CUL-2 is required for the G1-to-S-phase transition and mitotic chromosome condensation in Caenorhabditis elegans. *Nat Cell Biol* **1**, 486-492 (1999).
25. Wang, X.W. *et al.* GADD45 induction of a G2/M cell cycle checkpoint. *Proc Natl Acad Sci U S A* **96**, 3706-11 (1999).
26. Harkin, D.P. *et al.* Induction of GADD45 and JNK/SAPK-dependent apoptosis following inducible expression of BRCA1. *Cell* **97**, 575-86 (1999).
27. Moorman, J.P., Bobak, D.A. & Hahn, C.S. Inactivation of the small GTP binding protein Rho induces multinucleate cell formation and apoptosis in murine T lymphoma EL4. *J Immunol* **156**, 4146-53 (1996).
28. Shannon, K.B. & Li, R. The multiple roles of Cyk1p in the assembly and function of the actomyosin ring in budding yeast. *Mol Biol Cell* **10**, 283-96 (1999).
29. RayChaudhuri, D. ZipA is a MAP-Tau homolog and is essential for structural integrity of the cytokinetic FtsZ ring during bacterial cell division. *Embo J* **18**, 2372-83 (1999).
30. Radhakrishna, H., Al-Awar, O., Khachikian, Z. & Donaldson, J.G. ARF6 requirement for Rac ruffling suggests a role for membrane trafficking in cortical actin rearrangements. *J Cell Sci* **112**, 855-66 (1999).
31. Kondo, A. *et al.* A New Paxillin-binding Protein, PAG3/Papalpha/KIAA0400, Bearing an ADP-Ribosylation Factor GTPase-activating Protein Activity, Is Involved in Paxillin Recruitment to Focal Adhesions and Cell Migration. *Mol Biol Cell* **11**, 1315-1327 (2000).

32. Leeuwen, F.N. *et al.* The guanine nucleotide exchange factor Tiam1 affects neuronal morphology; opposing roles for the small GTPases Rac and Rho. *J Cell Biol* **139**, 797-807 (1997).
33. Arai, M. & Kwiatkowski, D.J. Differential developmentally regulated expression of gelsolin family members in the mouse. *Dev Dyn* **215**, 297-307 (1999).
34. Satoh, T., Kato, J., Nishida, K. & Kaziro, Y. Tyrosine phosphorylation of ACK in response to temperature shift-down, hyperosmotic shock, and epidermal growth factor stimulation. *FEBS Lett* **386**, 230-4 (1996).
35. Rohatgi, R. *et al.* The interaction between N-WASP and the Arp2/3 complex links Cdc42- dependent signals to actin assembly. *Cell* **97**, 221-31 (1999).
36. North, A.J. *et al.* Molecular map of the desmosomal plaque. *J Cell Sci* **112**, 4325-36 (1999).
37. Yamamoto, K., Ichijo, H. & Korsmeyer, S.J. BCL-2 is phosphorylated and inactivated by an ASK1/Jun N-terminal protein kinase pathway normally activated at G(2)/M. *Mol Cell Biol* **19**, 8469-78 (1999).
38. Phillips, A.C., Ernst, M.K., Bates, S., Rice, N.R. & Vousden, K.H. E2F-1 potentiates cell death by blocking antiapoptotic signaling pathways. *Mol Cell* **4**, 771-81 (1999).
39. Hakem, A., Sasaki, T., Kozieradzki, I. & Penninger, J.M. The cyclin-dependent kinase Cdk2 regulates thymocyte apoptosis. *J Exp Med* **189**, 957-68 (1999).
40. Gil-Gomez, G., Berns, A. & Brady, H.J. A link between cell cycle and cell death: Bax and Bcl-2 modulate Cdk2 activation during thymocyte apoptosis. *Embo J* **17**, 7209-18 (1998).
41. Lind, E.F. *et al.* Bcl-2-induced changes in E2F regulatory complexes reveal the potential for integrated cell cycle and cell death functions. *J Immunol* **162**, 5374-9 (1999).
42. Silva, M. *et al.* Erythropoietin Can Induce the Expression of Bcl-x(L) through Stat5 in Erythropoietin-dependent Progenitor Cell Lines. *J Biol Chem* **274**, 22165-22169 (1999).
43. Ross-Macdonald, P. *et al.* Large-scale analysis of the yeast genome by transposon tagging and gene disruption. *Nature* **402**, 413-418 (1999).
44. McNally, F.J. & Thomas, S. Katanin is responsible for the M-phase microtubule-severing activity in Xenopus eggs. *Mol Biol Cell* **9**, 1847-61 (1998).
45. Hartman, J.J. *et al.* Katanin, a microtubule-severing protein, is a novel AAA ATPase that targets to the centrosome using a WD40-containing subunit. *Cell* **93**, 277-87 (1998).
46. Farrance, I.K. & Ordahl, C.P. The role of transcription enhancer factor-1 (TEF-1) related proteins in the formation of M-CAT binding complexes in muscle and non-muscle tissues. *J Biol Chem* **271**, 8266-74 (1996).
47. Stewart, A.F., Suzow, J., Kubota, T., Ueyama, T. & Chen, H.H. Transcription factor RTEF-1 mediates alpha1-adrenergic reactivation of the fetal gene program in cardiac myocytes. *Circ Res* **83**, 43-9 (1998).
48. Lee, S.B. *et al.* The Wilms tumor suppressor WT1 encodes a transcriptional activator of amphiregulin. *Cell* **98**, 663-73 (1999).
49. White, K.P., Rifkin, S.A., Hurban, P. & Hogness, D.S. Microarray analysis of Drosophila Development During Metamorphosis. *Science* **286**, 2179-2184 (1999).
50. Martin, L.G., Demers, G.W. & Galloway, D.A. Disruption of the G1/S transition in human papillomavirus type 16 E7- expressing human cells is associated with altered regulation of cyclin E. *J Virol* **72**, 975-85 (1998).

(10) **Appendices**

None